TWO ERDŐS PROBLEMS ON LACUNARY SEQUENCES: CHROMATIC NUMBER AND DIOPHANTINE APPROXIMATION

YUVAL PERES AND WILHELM SCHLAG

Abstract. Let $\{n_k\}$ be an increasing lacunary sequence, *i.e.*, $n_{k+1}/n_k > 1+\epsilon$ for some $\epsilon > 0$. In 1987, P. Erdős asked for the chromatic number $\chi(G)$ of a graph G with vertex set $\mathbb{Z}$, where two integers $x, y \in \mathbb{Z}$ are connected by an edge iff their difference $|x-y|$ is in the sequence $\{n_k\}$. Y. Katznelson found a connection to a Diophantine approximation problem (also due to Erdős): the existence of $\theta \in (0,1)$ such that all the multiples $n_j\theta$ are at least distance $\delta(\theta) > 0$ from $\mathbb{Z}$. Katznelson showed that $\chi(G) \le C\epsilon^{-2}|\log\epsilon|$. We apply the Lovász local lemma to establish that $\delta(\theta) > c\epsilon|\log\epsilon|^{-1}$ for some θ, which implies that $\chi(G) < C\epsilon^{-1}|\log\epsilon|$. This is sharp up to the logarithmic factor.

1. **Introduction**

The chromatic number $\chi(\mathcal{G})$ of a graph $\mathcal{G}$ is the minimal number of colors that can be assigned to the vertices of $\mathcal{G}$ so that no edge connects two vertices of the same color. In 1987, Paul Erdős posed the following problem[1]:

Problem A: *Let $\epsilon > 0$ be fixed and suppose $\mathcal{S} = \{n_j\}_{j=1}^{\infty}$ is a sequence of positive integers such that $n_{j+1} > (1+\epsilon)n_j$ for all $j \ge 1$. Define a graph $\mathcal{G} = \mathcal{G}(\mathcal{S})$ with vertex set $\mathbb{Z}$ (the integers) by letting the pair (n,m) be an edge iff $|n-m| \in \mathcal{S}$. Is the chromatic number $\chi(\mathcal{G})$ finite?*

There is a related Diophantine approximation problem, posed earlier by Erdős [5]:

Problem B: *Let $\epsilon > 0$ and $\mathcal{S}$ be as above. Is there a number $\theta \in (0,1)$ so that the sequence $\{n_j\theta\}_{j=1}^{\infty}$ is not dense modulo 1?*

1991 *Mathematics Subject Classification.* Primary: 05C15 Secondary: 42A55, 11B05.

Key words and phrases. Lacunary sequence, chromatic number, diophantine approximation.

Research of Peres was partially supported by NSF grant DMS-0605166. Research of Schlag was partially supported by NSF grant DMS-0617854.

[1]according to Y. Katznelson [10].

The relation between Problem A and Problem B was discovered by Katznelson [10]: Let $\delta > 0$ and $\theta \in (0,1)$ be such that $\inf_j \|\theta n_j\| > \delta$, where $\|\cdot\|$ denotes the distance to the closest integer. Partition the circle $\mathbb{T} = [0,1)$ into $k = \lceil \delta^{-1} \rceil$ disjoint intervals $I_1, \ldots, I_k$ of length $\frac{1}{k} \leq \delta$. Let $\mathcal{G}$ be the graph from Problem A and color the vertex $n \in \mathbb{Z}$ with color j iff $n\theta \in I_j \pmod 1$. Clearly, any two vertices connected by an edge must have different colors. Therefore, $\chi(\mathcal{G}) \leq k = \lceil \delta^{-1} \rceil$. From this, one can easily deduce that $\chi(\mathcal{G}) < \infty$ for any lacunary sequence with ratio at least $1+\epsilon$ by partitioning into several subsequences (see the end of the introduction); however, the bound obtained this way grows exponentially in ϵ^{-1}.

Problem B was solved by de Mathan [11] and Pollington [15] and their proofs provide bounds on $\chi(\mathcal{G})$ that grow polynomially in ϵ^{-1}. More precisely, they show that there exists $\theta \in (0,1)$ such that

$$\inf_{j \geq 1} \|\theta n_j\| > c\epsilon^4 |\log \epsilon|^{-1}$$

where $c > 0$ is some constant. This bound was improved by Katznelson [10], who showed that there exists a θ such that

$$\inf_{j \geq 1} \|\theta n_j\| > c\epsilon^2 |\log \epsilon|^{-1}. \tag{1.1}$$

Akhunzhanov and Moshchevitin [1] removed the logarithmic factor on the right hand side of (1.1), see also Dubickas [4].

We can now state the main result of this note.

Theorem 1.1. *Suppose $\mathcal{S} = \{n_j\}$ satisfies $n_{j+1}/n_j \geq 1+\epsilon$, where $0 < \epsilon < 1/4$. Then there exists $\theta \in (0,1)$ such that*

$$\inf_{j \geq 1} \|\theta n_j\| > c\epsilon |\log \epsilon|^{-1}\,, \tag{1.2}$$

where $c > 0$ is a universal constant. Therefore, the graph $\mathcal{G} = \mathcal{G}(\mathcal{S})$ described in Problem A satisfies $\chi(\mathcal{G}) \leq 1 + c^{-1}\epsilon^{-1}|\log \epsilon|$.

Up to the $|\log \epsilon|^{-1}$ factor, (1.2) cannot be improved. Indeed, let $n_j = j$ for $j = 1, 2, \ldots, \lfloor \epsilon^{-1} \rfloor$ and continue this as a lacunary sequence with ratio $1+\epsilon$. It is clear that $\chi(\mathcal{G}) > \lfloor \epsilon^{-1} \rfloor$ in this case, so that the power of ϵ in (1.2) cannot be decreased.

In order to prove Theorem 1.1, we use the Lovász local lemma from probabilistic combinatorics, see [6] or [2, Chap. 5]. Loosely speaking, given events $A_1, A_2, \ldots$ in a probability space, this lemma bounds $\mathbb{P}\Big(\bigcap_{j=1}^N A_j^c\Big)$ from below if the events A_j have small probability and

each A_i is almost independent of most of the others. See the following section for a precise statement. Theorem 1.1 is established in Section 3.

Finally, we recall how the aforementioned relation between problems A and B yields an easy proof that $\chi(\mathcal{G}) < \infty$ for any $\epsilon > 0$. See [10], [19, Chap. 5] and [17] for variants of this argument. First suppose that $n_{j+1}/n_j > 4$ for all j. In this case,

$$\bigcap_{j=1}^{\infty} \Big\{ \theta \in \mathbb{T} \,:\, \|\theta n_j\| > \frac{1}{4} \Big\} \neq \emptyset. \tag{1.3}$$

Indeed, fix some j and notice that the set $\{\theta \in \mathbb{T} \,:\, \|\theta n_j\| > \frac{1}{4}\}$ is the union of the middle halves of intervals $[\frac{\ell}{n_j}, \frac{\ell+1}{n_j}]$ where $\ell = 0, 1, \ldots, n_j - 1$. Since $n_{j+1} > 4n_j$, each such middle half contains an entire interval of the form $[\frac{\ell'}{n_{j+1}}, \frac{\ell'+1}{n_{j+1}}]$. Iterating this yields a sequence of nested intervals and establishes (1.3).

Now suppose just that $n_{j+1}/n_j > 1 + \epsilon > 1$. Pick $K = \lceil 2\epsilon^{-1} \rceil$ so that $(1+\epsilon)^K > 4$. Divide the given sequence $\mathcal{S}$ into K subsequences $\{n_{Kj+r}\}_{j=0}^{\infty}$, with $r = 1, 2, \ldots, K$. Applying (1.3) to each such subsequence yields $(\theta_1, \ldots, \theta_K) \in \mathbb{T}^K$ so that

$$\inf_{j \geq 0} \|n_{Kj+r}\, \theta_r\| \geq \frac{1}{4} \;\text{ for all } r = 1, 2, \ldots, K.$$

Coloring each integer m according to which quarter of the unit interval $m\theta_r$ falls into for $1 \leq r \leq K$, shows that $\chi(\mathcal{G}) \leq 4^K$. Observe that as $\epsilon \to 0$, this bound grows exponentially in $\frac{1}{\epsilon}$.

2. **A one-sided version of the local lemma**

The following lemma is the variant of the Lovász local lemma [6] that we apply to Problem B above. Since it is not exactly stated in this form (neither in terms of the hypotheses nor the conclusion) in [6] or [2], we provide a proof for the reader's convenience. We stress, however, that it is a simple adaptation of the argument given in chapter 5 of [2].

Lemma 2.1. *Let $\{A_j\}_{j=1}^N$ be events in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let $\{x_j\}_{j=1}^N$ be a sequence of numbers in $(0,1)$. Assume that for every $i \leq N$, there is an integer $0 \leq m(i) < i$ so that*

$$\mathbb{P}\Big(A_i \Big| \bigcap_{j<m(i)} A_j^c \Big) \;\leq\; x_i \prod_{j=m(i)}^{i-1} (1 - x_j). \tag{2.1}$$

Then for any integer $n \in [1, N]$, we have

$$\mathbb{P}\Big(\bigcap_{i=1}^{n} A_i^c\Big) \;\; \geq \;\; \prod_{\ell=1}^{n}(1 - x_\ell). \tag{2.2}$$

Proof. Denote $B_1 = \Omega$ and $B_\ell = A_1^c \cap \ldots \cap A_{\ell-1}^c$ for $\ell > 1$. We claim that for each $\ell \geq 1$,

$$\mathbb{P}(A_\ell \,|B_\ell) \;\; \leq \;\; x_\ell. \tag{2.3}$$

Since

$$\mathbb{P}\Big(\bigcap_{i=1}^{n} A_i^c\Big) = \prod_{\ell=1}^{n}\Big(1 - \mathbb{P}(A_\ell \,|B_\ell)\Big),$$

(2.3) implies (2.2). The claim (2.3) is verified inductively:

$$\mathbb{P}\Big(A_\ell \,\Big|B_\ell\Big) = \frac{\mathbb{P}\Big(A_\ell \cap B_\ell \,\Big|B_{m(\ell)}\Big)}{\mathbb{P}\Big(B_\ell \,\Big|B_{m(\ell)}\Big)} \leq \frac{\mathbb{P}\Big(A_\ell \,\Big|B_{m(\ell)}\Big)}{\mathbb{P}\Big(B_\ell \,\Big|B_{m(\ell)}\Big)}\,. \tag{2.4}$$

The denominator in the rightmost fraction can be written as a product

$$\mathbb{P}\Big(B_\ell \,\Big|B_{m(\ell)}\Big) = \prod_{j=m(\ell)}^{\ell-1}\Big(1 - \mathbb{P}(A_j \,|B_j)\Big)\,. \tag{2.5}$$

By the inductive hypothesis, this product is at least $\prod_{j=m(\ell)}^{\ell-1}(1-x_j)$, whereas the numerator in the right-hand side of (2.4) is at most $x_\ell \prod_{j=m(\ell)}^{\ell-1}(1-x_j)$ by (2.1). This finishes the proof. $\square$

3. **Rotation orbits sampled along a lacunary sequence**

In this section we present our quantitative result on problem B, which extends Theorem 1.1 and is also applicable to unions of lacunary sequences. Observe that if $\mathcal{S} = \{n_j\}$ is lacunary with ratio $1 + \epsilon$, then it satisfies the hypothesis of the next theorem with $M = \lceil \epsilon^{-1} \rceil$. More generally, if $\mathcal{S}_1, \ldots, \mathcal{S}_\ell$ are lacunary with ratios $1 + \epsilon_1, \ldots, 1 + \epsilon_\ell$ respectively, then their union satisfies the hypothesis with $M = \sum_{i=1}^{\ell} \lceil \epsilon_i^{-1} \rceil$. We shall assume that $M \geq 4$.

Theorem 3.1. *Suppose $\mathcal{S} = \{n_j\}$ satisfies $n_{j+M} > 2n_j$ for all j. Define*

$$E_j \;\; = \;\; \left\{\theta \in \mathbb{T} \,:\, \|n_j\theta\| < \frac{c_0}{M \log_2 M}\right\} \tag{3.1}$$

for $j \geq 1$. If $240\, c_0 \leq 1$, then

$$\bigcap_{j=1}^{\infty} E_j^c \;\; \neq \;\; \emptyset. \tag{3.2}$$

Proof. Set

$$\delta \;\; = \;\; \frac{c_0}{M \log_2 M}. \tag{3.3}$$

For each $j = 1, 2, \ldots$ define an integer ℓ_j by

$$2^{-\ell_j - 1} < \frac{2\delta}{n_j} \leq 2^{-\ell_j}. \tag{3.4}$$

Let A_j be the union of all the open dyadic intervals of size $2^{-\ell_j}$ that intersect E_j. Observe that E_j is the union of n_j intervals of length $\frac{2\delta}{n_j}$, and each one of them is covered by at most two dyadic intervals of length $2^{-\ell_j}$. Therefore,

$$\mathbb{P}(A_j) \leq 2 \cdot 2^{-\ell_j} n_j \leq 8\delta. \tag{3.5}$$

where $\mathbb{P}$ is Lebesgue measure on $[0, 1]$. Define

$$h \;\; = \;\; \lceil C_1 \log_2 M \rceil M, \tag{3.6}$$

where $C_1 \geq 5$ is a constant to be determined. Our goal is to apply Lemma 2.1 with $m(i) = i - h$ and $x_i = x = h^{-1}$ for all $i \in \mathbb{Z}^+$. To verify (2.1), fix some $i > h$. Then

$$\bigcap_{j<i-h} A_j^c = \bigcup_s I_s$$

with dyadic intervals I_s of length $|I_s| = 2^{-\ell_{i-h-1}}$. Hence

$$\begin{aligned}
&\mathbb{P}\Big(A_i \cap \bigcap_{j<i-h} A_j^c\Big) = \sum_s \mathbb{P}(A_i \cap I_s) \leq \sum_s (1 + |I_s| n_i) 2^{1-\ell_i} \\
&\leq \mathbb{P}\Big(\bigcap_{j<i-h} A_j^c\Big)\Big[2\, n_i 2^{-\ell_i} + 2^{1+\ell_{i-h-1}-\ell_i}\Big] \\
&\leq \mathbb{P}\Big(\bigcap_{j<i-h} A_j^c\Big)\Big[8\delta + 4\frac{n_{i-h-1}}{n_i}\Big].
\end{aligned} \tag{3.7}$$

To pass to (3.7), one uses (3.4). By (3.6),

$$\frac{n_{i-h-1}}{n_i} \leq 2^{-C_1 \log_2 M} = M^{-C_1}. \tag{3.8}$$

Inserting this bound into (3.7) yields

$$\mathbb{P}\Big(A_i \Big| \bigcap_{j<i-h} A_j^c\Big) \le 12\delta, \quad \text{provided that} \tag{3.9}$$

$$\frac{c_0 M^{C_1}}{M\log_2 M} \ge 1, \ \text{ which is the same as } \ M^{-C_1} \le \delta. \tag{3.10}$$

By (3.5), the estimate (3.9) holds also if $i \le h$. In order to satisfy (2.1), we need to ensure that

$$12\delta \le x(1-x)^h. \tag{3.11}$$

Since $x = h^{-1} \le 1/16$, we have $(1-x)^h \ge 1/3$. Thus (3.11) will be satisfied if

$$36\delta \le x = h^{-1}. \tag{3.12}$$

By (3.3) and (3.6), $h\delta = \lceil C_1 \log_2 M \rceil\, c_0 / \log_2 M \le \frac{10}{9} C_1 c_0$, since $M \ge 4$ and $C_1 \ge 5$. Therefore, (3.12) will hold if

$$40 C_1 c_0 \le 1\,. \tag{3.13}$$

Take $C_1 = 6$ and $c_0 = \frac{1}{240}$; then (3.13) and (3.10) are both satisfied. By Lemma 2.1,

$$\mathbb{P}\Big(\bigcap_{i=1}^n A_i^c\Big) \ \ge \ \Big(1-x\Big)^n. \tag{3.14}$$

Since each of the A_j^c is compact, and $A_j^c \subset E_j^c$, the intersection $\cap_{j=1}^\infty E_j^c$ is nonempty, as claimed. $\square$

4. **Intersective sets**

Let $\mathcal{S} = \{n_j\}$ with $n_{j+1} \ge (1+\epsilon)n_j$ be a lacunary sequence of positive integers with ratio $1+\epsilon > 1$. Denote $M = \lceil \epsilon^{-1} \rceil$. We have shown that there is a coloring of the graph $\mathcal{G}$ with at most $CM\log M$ colors. Let $\mathcal{A}_{\max}$ be a set of integers of the same color with upper density

$$D^*(\mathcal{A}_{max}) = \limsup_{N\to\infty} \frac{\operatorname{card}(\mathcal{A}_{\max} \cap [-N,N])}{2N+1} > \frac{c}{M\log M},$$

where c is a constant. By the definition of coloring,

$$(\mathcal{A}_{\max} - \mathcal{A}_{\max}) \cap \mathcal{S} = \emptyset.$$

A set $\mathcal{H}$ is called *intersective* if

$$(\mathcal{A} - \mathcal{A}) \cap \mathcal{H} \ne \emptyset$$

for any $\mathcal{A} \subset \mathbb{Z}$ with $D^*(\mathcal{A}) > 0$. Intersective sets are precisely the Poincaré sets considered by Fürstenberg [8], see [3] and [13]. Generally speaking, it is not a simple matter to decide whether a given set is intersective or not. Fürstenberg [7] and Sárközy [18] showed that the squares (and more generally, the set $\{P(n)\}_{n\in\mathbb{Z}}$ where P is a polynomial over $\mathbb{Z}$ that vanishes at some integer) are intersective. In particular, intersective sets can have zero density. On the other hand, the previous discussion shows that any finite union of lacunary sequences is not intersective. There are some related concepts of intersectivity which we briefly recall; for a nice introduction to this subject see chapter 2 in Montgomery's book [12]. A set of integers $\mathcal{H}$ is called a *van der Corput* set, if for any sequence $\{x_j\}_j$ of numbers with the property that $\{x_{j+h} - x_j\}_j$ is uniformly distributed modulo 1 for all $h \in \mathcal{H}$, one has that $\{x_j\}_j$ is uniformly distributed modulo 1. It was shown by Kamae and Mendes–France [9] that any van der Corput set is intersective. There is a more quantitative version of this fact, due to Ruzsa [16]: Define

$$\gamma_{\mathcal{H}} = \inf \int_0^1 T(x)\,dx \quad \text{where} \quad T(x) = a_0 + \sum_{h\in\mathcal{H}} a_h \cos(2\pi hx) \geq 0 \quad \text{and} \quad T(0) = 1,$$

T being a trigonometric polynomial. It it is known that $\mathcal{H}$ is a van der Corput set iff $\gamma_{\mathcal{H}} = 0$, see [16] and [12]. Also, let

$$\delta_{\mathcal{H}} = \sup\{D^*(\mathcal{A}) \,:\, (\mathcal{A} - \mathcal{A}) \cap \mathcal{H} = \emptyset\}.$$

By definition, $\mathcal{H}$ is intersective iff $\delta_{\mathcal{H}} = 0$. It was shown in [16] that $\delta_{\mathcal{H}} \leq \gamma_{\mathcal{H}}$. In view of this fact, our Theorem 1.1 has the following consequence.

Corollary 4.1. *Let $\mathcal{S} = \{n_j\}_j$ be lacunary with ratio $1+\epsilon$, and suppose T is a nonnegative trigonometric polynomial of the form*

$$T(x) = a_0 + \sum_j a_j \cos(2\pi n_j x) \quad \textit{with } T(0) = 1.$$

Then for some universal constant $c > 0$,

$$\int_0^1 T(x)\,dx \;=\; a_0 > c\epsilon\,|\log\epsilon|^{-1}. \tag{4.1}$$

If one could show that the bound given by (4.1) is optimal, then it would follow that the $|\log\epsilon|^{-1}$ factor in (1.2) cannot be removed.

Remark. The first author heard Y. Katznelson present his proof of (1.1) in a lecture at Stanford in 1991, but that proof only appeared ten years later [10]. The proof of our main result, Theorem 1.1, was obtained in 1999 and presented in a lecture [14] at the IAS, Princeton in 2000. We thank J. Grytczuk for urging us to publish this result and providing references for recent work on related problems.

References

[1] Akhunzhanov, R.K. and Moshchevitin, N. G., *On the Chromatic Number of the Distance Graph Associated with a Lacunary Sequence.* Doklady Akademii Nauk. Ross. **397** (2004), 295–296.

[2] Alon, N., Spencer, J. *The probabilistic method.* Wiley and Sons, New York, 1992.

[3] Bertrand-Mathis, A. *Ensembles intersectifs et rcurrence de Poincaré.* Colloque de Théorie Analytique des Nombres "Jean Coquet" (Marseille, 1985), 55–72, Publ. Math. Orsay, 88-02, Univ. Paris XI, Orsay, 1988.

[4] Dubickas, A. On the fractional parts of lacunary sequences. Mathematica Scand. **99** (2006), 136–146.

[5] Erdős, P. *Repartition mod* 1. Lecture Notes in Math. 475, Springer–Verlag, N.Y., 1975.

[6] Erdős, P., Lovász, L. *Problems and results on 3–chromatic hypergraphs and some related problems.* In A. Hajnal et al., eds., Infinite and Finite Sets, North–Holland, Amsterdam, (1975), 609–628.

[7] Fürstenberg, H. *Ergodic behavior of diagonal measures and a theorem of Szemerdi on arithmetic progressions.* J. Analyse Math. **31** (1977), 204–256.

[8] Fürstenberg, H. *Poincaré recurrence and number theory.* Bull. Amer. Math. Soc. (N.S.) **5** (1981), 211–234.

[9] Kamae, T., Mendes–France, M. *van der Corput's difference theorem.* Israel Journal Math. **31** (1978), 335–342.

[10] Katznelson, Y. *Chromatic numbers of Cayley graphs on $\mathbb{Z}$ and recurrence.* Combinatorica 21 (2001), 211–219.

[11] de Mathan, B. *Numbers contravening a condition in density modulo* 1. Acta Math. Acad. Sci. Hungar. **36** (1980), 237–241.

[12] Montgomery, H. *Ten lectures on the interface between analytic number theory and harmonic analyis.* CBMS lectures notes #84, AMS 1994.

[13] Peres, Y. *Application of Banach limits to the study of sets of integers.* Israel J. Math. **62** (1988), 17–31.

[14] Peres, Y. *Two Erdos problems on lacunary sequences: Chromatic number and diophantine approximation.* Lecture at the Institute of Advanced Study, Princeton, New Jersey, January 2000. `http://www.math.ias.edu~warfield/abstr2000.html#peres`

[15] Pollington, A. D. *On the density of the sequence* $\{n_k\xi\}$. Illinois J. Math. **23** (1979), no. 4, 511–515.

[16] Ruzsa, I. Z. *Connections between the uniform distribution of a sequence and its differences*, Topics in Classical Number Theory (Budapest 1981), North-Holland, Amsterdam (1984), 1419–1443.

[17] Ruzsa, I. Z., Tuza, Z. and Voigt, M. *Distance Graphs with Finite Chromatic Number*. Jour. Combin. Theory B **85** (2002), 181–187.

[18] Sárközy, A. *On difference sets of sequences of integers. I.* Acta Math. Acad. Sci. Hungar. **31** (1978), 125–149.

[19] Weiss, B. *Single orbit dynamics*. CBMS Regional Conference Series in Mathematics, **95**. Amer. Math. Soc., Providence, RI (2000).

Microsoft Research, Redmond, WA and University of California, Berkeley, CA
E-mail address: peres@microsoft.com

Department of Mathematics, University of Chicago.
E-mail address: schlag@math.uchicago.edu